

Unscharfe Klassifikation von Webkennzahlen

Das Customer Relationship Management braucht Methoden und Kennzahlen, um Website-Besucher und Onlinekunden sowie deren Wert für das eBusiness zu klassifizieren und zu bewerten.

In der Praxis weitverbreitete Kundenbewertungs- und Segmentierungsmethoden sind Scoring-Modelle. Ein Problem solcher Heuristiken besteht darin, dass Kundenstrategien anhand von trennscharfen Klassen und Scores definiert werden. Aufgrund scharfer Klassen können Klassifikationsfehler auftreten, Kunden falsch bewertet und das Marketingbudget bei der Kundenansprache oder beim Kundenmanagement suboptimal eingesetzt werden.

Dieser Beitrag schlägt eine unscharfe Scoring-Methode vor, konkret ein unscharfes eRFM-Modell. Die Werte der Variablen Recency, Frequency und Monetary Value sind einfach zu erfassen und dank der unscharfen Klassifikation genauer zu segmentieren als bei trennscharfer Klassifikation.

Dies wird anhand einer Simulation und einer Fallstudie eines Unternehmens im Online-Detailhandel gezeigt.

Inhaltsübersicht

- 1 Erfolgskette des Onlinemarketings
- 2 eRFM: Einfach messbare und aussagekräftige eBusiness-Kennzahlen
- 3 Klassifikationsmethoden
 - 3.1 Scharfe Scoring-Modelle und Problematik
 - 3.2 Der Ansatz der unscharfen Klassifikation
 - 3.3 Unscharfe Klassifikation von Frequency und Monetary Value
 - 3.4 Fuzzy eRFM Data Cube

- 4 Fallstudie coop@home
 - 4.1 Daten der Fallstudie
 - 4.2 Berechnung der Fuzzy-Werte
 - 4.3 Vergleich scharfe und unscharfe Klassifikation
- 5 Unscharfe Klassifikation weiterer Web- und eCommerce-Kennzahlen
- 6 Nutzenbetrachtung
- 7 Literatur

1 Erfolgskette des Onlinemarketings

Onlinekunden sind eine der wichtigsten Ressourcen eines jeden eCommerce-Unternehmens. Da der Onlinekunde persönlich meist nicht bekannt ist, braucht es Instrumente und Techniken wie Web Analytics, die das Verhalten, aber auch die Bedürfnisse und Wünsche der Kunden eruieren. *Web Analytics*, im Rahmen des Web Mining auch *Web Usage Mining* [Liu 2007] genannt, ist die Messung, Analyse und Auswertung von Webmetriken zwecks Website-Optimierung.

Voraussetzung jedes eBusiness ist eine funktionierende Infrastruktur, wie Server, Content-Management- und eShop-Systeme, aber auch Mitarbeiter, Produktion und Distribution (❶ in Abb. 1). Die im eShop angebotenen Produkte und Dienstleistungen müssen die Erwartungen der Kunden erfüllen und einen spürbaren Nutzen generieren. Der *Kundennutzen*, der durch die Website und das Onlineangebot generiert wird, trägt zur *Zufriedenheit* von Onlinekunden bei (❷). Auch die wahrgenommene *Website-Qualität*, die *Website-Zufriedenheit* und die *Gestaltung der Website*, wie z.B. die Struktur, das Design, der Informationswert, die Funktionalität und der Unterhaltungswert, steuern maßgeblich zum Erfolg eines Internetauftritts bei [Wünschmann et al. 2008].

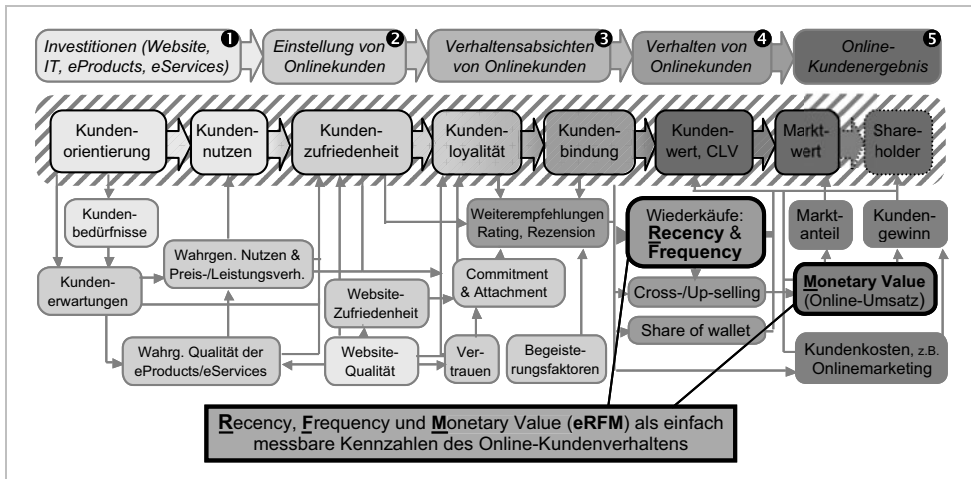


Abb. 1: Erfolgskette des Onlinemarketings

Die CRM-Literatur wie [Bruhn & Homburg 2008] untermauert, dass zunehmende Kundenzufriedenheit auf der Ebene der Verhaltensabsicht (3) die Loyalität und Bindung von On-/Offlinekunden erhöht. Eine hohe Online-Kundenloyalität, sprich der beabsichtigte oder tatsächliche wiederholte Besuch und Einkauf auf der Website, ist schwieriger zu erreichen und als wichtiger einzustufen als im Offlinegeschäft. Denn die Konkurrenz und ihr Angebot sind nur wenige Klicks entfernt. Immerhin ist die Loyalität online leichter zu messen, da das *Klickverhalten* (4) von Usern auf der Website genau ausgewertet werden kann. Die Auswertung geschieht entweder durch die Analyse der *Logfiles* oder durch *clientseitige Datensammlung* (d.h. mittels Web-Analytics-Software wie z.B. Google Analytics). Web Usage Mining und Web Analytics helfen, das Besucherverhalten zu verstehen und die Website zu optimieren (vgl. [Hassler 2008]).

Eine hohe Kundenzufriedenheit und Kundenbindung ermöglicht zukünftige Einnahmen und senkt die Transaktionskosten sowie die Kosten der Onlinekommunikation und des Verkaufs. In der Folge steigen tendenziell die *Online-Umsätze* und der Cashflow (das Online-Kun-

denergebnis 5). Gleichzeitig bedeuten eine stabile Kundenbasis aufgrund von Wiederkäufen sowie das durch *Empfehlungssysteme* forcierte Cross- und Up-Selling (d.h. der Verkauf von weiteren und höherwertigen Produkten) eine voraussehbare Quelle zukünftiger Umsätze.

Virales Marketing, Mund-zu-Mundpropaganda, Weiterempfehlungen, positive Bewertungen und Rezensionen zufriedener oder gar begeisterter Onlinekunden führen zu tieferen Akquisitions- und Onlinemarketing-Kosten sowie zu zusätzlichen Umsätzen von neuen Usern und Onlinekunden. Deshalb wirkt sich die Website- und Kundenzufriedenheit positiv auf den *Unternehmensgewinn* und auf den *Marktwert* des Unternehmens aus, indem die Volatilität und das Risiko antizipierter Cashflows reduziert wird [Anderson et al. 2004].

2 eRFM: Einfach messbare und aussagekräftige eBusiness-Kennzahlen

Während die Einstellungen und Verhaltensabsichten von Onlinekunden schwierig zu erfassen sind (z.B. mittels Onlinebefragung), ist das tatsächliche Verhalten auf der Website und vor

allem das vorliegende Kundenergebnis relativ einfach zu messen. Wann und wie oft ein Besucher auf der Website war (Anzahl Besuche bzw. Visits) und über welche Produkte er sich im eShop informiert, ist der Web-Analytics-Software zu entnehmen. Welche Produkte zu welchem Wert zu welchem Zeitpunkt ein Onlinekunde gekauft hat, ist unter anderem anhand der *Kundendatenbank* in Erfahrung zu bringen.

Mit *eRFM* schlägt dieser Beitrag drei Kennzahlen vor, mit denen der Wert eines Onlinekunden eingeschätzt werden kann (vgl. Tab. 1):

- **Recency**: die Zeitdauer seit dem letzten Online-Einkauf eines Kunden
- **Frequency**: die Anzahl der Online-Einkäufe
- **Monetary Value**: der monetäre Wert eines Kunden gemessen anhand des Online-Umsatzes

Das RFM-Modell ist ein bekanntes Kundenbewertungsmodell, wurde bisher aber nicht auf das Onlinemarketing angewandt. Zudem werden die drei Kennzahlen typischerweise trennscharf klassifiziert.

Empirisch sind eine hohe Kauffrequenz und ein hoher monetärer Wert von Kunden in Zukunft umso *wahrscheinlicher* (siehe z.B. [Krafft 2007]), je

- *kürzer* der letzte Einkauf zurückliegt (Recency), je
- *öfters* ein Kunde in einer bestimmten Periode eingekauft hat (Frequency) und je
- *höher* die Kundenumsätze in den letzten drei Jahren waren (Monetary Value).

Verfügt eine Website über keinen elektronischen Shop, so wird anstelle der Einkaufsfrequenz die Besuchsfrequenz betrachtet. Anstelle des Umsatzes können bei (Non-)Monetary Value auch andere Größen evaluiert werden, die der Website oder dem Unternehmen einen Mehrwert generieren, z.B. die Klickrate auf bezahlte Werbebanner, die Anzahl Downloads von Dokumenten, Ratings, die Anzahl und Qualität von User Generated Content, Kontakte, Anfragen oder die Newsletter-Einschreibung.

	Variable	Definition	Messeinheit
<i>Website mit eShop</i>	Recency	Die Kaufsaktualität steht für die Kürze (verstrichene Zeit) seit dem letzten Online-Einkauf eines Kunden	In Stunden, Tagen oder Wochen
	Frequency	Die Kaufsfrequenz (Anzahl Einkäufe) steht für die Häufigkeit, mit der ein Onlinekunde im eShop einkauft	Absolute/relative Anzahl Käufe
	Monetary Value	Der monetäre Wert eines Kunden wird gemessen anhand des Online-Umsatzes bzw. -Deckungsbeitrages	In Euro
<i>Website ohne eShop</i>	Recency	Die Besuchsaktualität steht für die Kürze (verstrichene Zeit) seit dem letzten Website-Besuch eines Besuchers	In Stunden, Tagen oder Wochen
	Frequency	Die Besuchsfrequenz (Anzahl Besuche) definiert, wie oft ein Besucher in einer Periode die Website besucht	Anzahl Besuche; Besuchsdauer
	Monetary/Non-Monetary Value	Aktivitäten wie z.B. Klickrate (auf Banner), Downloads, Ratings, Rezensionen, Newsletter-Einschreibung	In Euro; Index; Bewertung; Rating

Tab. 1: Definition von eRFM

3 Klassifikationsmethoden

3.1 Scharfe Scoring-Modelle und Problematik

Der skizzierte RFM-Ansatz kann den sogenannten *Scoring-Modellen* zugeordnet werden. Scoring-Modelle sind in der Praxis neben der ABC- und Portfolioanalyse die meistangewandte Klassifikationsmethode der *Kundenbewertung* und *-segmentierung* [Neckel & Knobloch 2005; Günter & Helm 2006; McCarthy & Hastak 2006]. Bei den in der Literatur diskutierten und in der Praxis angewandten Scoring-Modellen erfolgen die Klassifikationen von Variablen und die Verteilung von Punkten immer *trennscharf*.

Auch die Zuordnung der Punktwerte (als Kundenwertmaß) zu jeder Klasse im Beispiel in Tabelle 2 geschieht in einer trennscharfen Art und Weise: Wenn ein Kunde z.B. für 99 € einkauft, wird er der *Monetary-Klasse* »51-100 €« zugeordnet und erhält +25 Punkte. Hätte er nur um 2 € mehr eingekauft, wäre der Kunde in einer höheren Klasse klassifiziert und erhielte +35 Punkte.

Eine solche scharfe Klassifikation ist aber *problematisch*, wie das folgende Beispiel illus-

triert: Die fiktiven Onlinekunden Weber und Müller haben bei vier Variablen die gleichen Werte (vgl. Tab. 2 rechts); beide haben den gleichen Startwert (+25), die gleiche Anzahl Bestellungen (+42), Anzahl Retouren (0) und eMails (-2). Es gibt nur zwei kleine Unterschiede: Der letzte Einkauf von Weber liegt 28 Wochen zurück (Klasse »7-9 Monate«), und Müllers letzter Einkauf war vor 27 Wochen (Klasse »6 Monate«). Der durchschnittliche Umsatz von Müller (102 €; Klasse: »101-150 €«) war 3 € höher als jener von Weber (99 €; Klasse: »51-100 €«).

Obwohl Müller nur eine Woche früher als Weber einkaufte und gerade mal einen 3 € höheren Umsatz hatte, sind die beiden unterschiedlichen Klassen zugeordnet, und Müller erhält mit 140 Punkten 25 Punkte mehr als Weber mit 115. In anderen Worten: Obwohl Müller und Weber *dasselbe* Kaufverhalten hatten, erhielt Müller 20 % *mehr* Punkte als Weber und wird in der Folge im Rahmen des Marketings oder Kundenmanagements vom Unternehmen womöglich anders betreut.

Auch der umgekehrte Fall kann auftreten: Obwohl zwei Kunden völlig *unterschiedliches*

Faktoren	Punktwerte pro Faktor						Onlinekunden	
							Weber	Müller
Startwert	+25						+25	+25
Kaufaktualität (Recency)	6 Monate +40	7-9 Monate +25	10-12 Monate +15	13-18 Monate +5	19-24 Monate -5	Früher: -15	28 Wochen +25	27 Wochen +40
Anzahl Bestellungen (Frequency)	Zahl der Bestellungen multipliziert mit 6						7·6 = +42	7·6 = +42
Ø-Umsatz der letzten 3 Käufe (Monetary)	25 €: +5	26-50 €: +15	51-100 €: +25	101-150 €: +35	151-200 €: +40	>200 €: +45	99 €: +25	102 €: +35
Anzahl Retouren (kumuliert)	0-1: 0	2-3: -5	4-6: -10	7-10: -20	11-15: -30	>15: -40	0: 0	0: 0
Anzahl Werbesendungen	Je Hauptkatalog: -12		Je Sonderkatalog: -6		Je E-Mail: -2		1 E-Mail: -2	1 E-Mail: -2
Punktwert Total							115	140

Quelle: in Anlehnung an [Neckel & Knobloch 2005, S. 211]

Tab. 2: Klassisches Beispiel der RFM-Methode mit scharfen Klassen

Kaufverhalten an den Tag legen, erzielen sie *gleich viele* Punkte.

Solche *Fehlklassifikationen* sind typische Folgen von Heuristiken mit scharfen Klassen, die in der Praxis aufgrund der einfachen Handhabung in Kauf genommen werden. Erstaunlich ist jedoch, dass die Ungenauigkeit und Problematik von scharfen Klassifikationsmethoden in der wissenschaftlichen Literatur kaum kritisiert werden, obwohl sie gerade im Kundenmanagement weitreichende Konsequenzen auf das Geschäft haben können. Denn einerseits werden *Kundenpotenziale nicht ausgeschöpft*, wenn eigentlich »gute«, wertvolle Kunden fälschlicherweise als »schlecht« klassifiziert werden und unterbewertet sind. Andererseits werden im Marketing *Ressourcen verschwendet*, wenn eher wenig wertvolle Kunden irrtümlich als privilegierte und damit kostspielige Key Accounts eingestuft werden. Mit dem Ansatz der unscharfen Klassifikation können solche Probleme reduziert werden.

3.2 Der Ansatz der unscharfen Klassifikation

Die Theorie der *unscharfen Logik* und der unscharfen Mengen geht auf Lofti A. Zadeh und das Jahr 1965 zurück [Zadeh 1965]. Beim Ansatz mit scharfen Mengen wird ein Objekt immer genau einer Menge (Klasse) zugeordnet. Klassifiziert man zum Beispiel die Kauffrequenz, so wird die *Anzahl Online-Einkäufe* von 0 bis 16 Einkäufen als »tiefe« (Klasse C1), von 17 bis 33 als »mittlere« (C2) und bei mehr als 34 Einkäufen als »hohe Anzahl« bewertet (C3; vgl. Abb. 2a).

In einer scharfen Klassifikation gehören Kunde 1 mit 32 Einkäufen und Kunde 2 mit 35 trotz ähnlicher Einkaufsfrequenz zu zwei verschiedenen Klassen (Kunde 1 hat eine »mittlere« und Kunde 2 eine »hohe Anzahl« Einkäufe). Kunde 1 wird gegenüber Kunde 2 diskriminiert: Hätte er nur zwei Einkäufe mehr getätigt, wäre er auch in die Klasse »hohe Anzahl« gerutscht. In einer unscharfen Klassifikation mit kontinuierlichem Klassenübergang (Abb. 2b) gehören beide Kunden zu *zwei Klassen gleichzeitig*: Kunde 1 mit 0,55 (55 %) zur Klasse »mittlere« und mit 0,45 (45 %) zur Klasse »hohe Anzahl« Einkäufe; Kunde 2 mit 40 % zur Klasse »mittlere« und mit 60 % zur Klasse »hohe Anzahl« Einkäufe. Eine unscharfe Klassifikation ist also fairer als eine scharfe, und den einzelnen Werten der klassifizierten Objekte wird Rechnung getragen.

3.3 Unscharfe Klassifikation von Frequency und Monetary Value

Um unscharfe Klassen aus den scharfen Klassen und Kontexten herzuleiten, werden den Äquivalenzklassen als Erstes *linguistische Variablen* und *Terme* zugeordnet [Zimmermann 1992]. Mithilfe von verbalen Begriffen (d.h. Wörtern oder Wortkombinationen) können Äquivalenzklassen intuitiv formuliert werden, was dem menschlichen Verständnis- und Urteilsvermögen entgegenkommt.

Im folgenden Beispiel wird der verbale Begriff Kauffrequenz (Frequency) durch die Terme »tief« und »hoch« beschrieben. Der Online-Umsatz besitzt die Terme »schwach« und »stark« (vgl. Abb. 3).

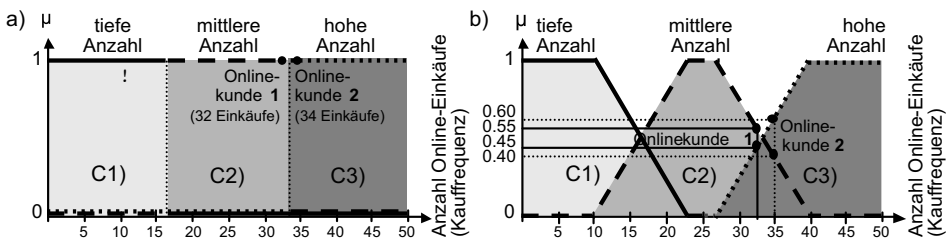


Abb. 2: Scharfe (a) und unscharfe (b) Klassifikation der Kauffrequenz

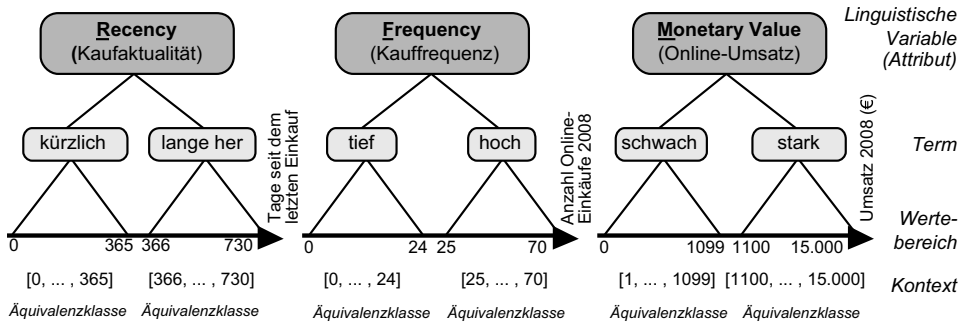


Abb. 3: Das Konzept der linguistischen Variablen Recency, Frequency und Monetary Value (RFM)

Die Kontexte und Wertebereiche werden wie folgt definiert:

- *Frequency*, gemessen anhand der Anzahl Einkäufe, hat den Wertebereich [0, 70] und zwei Äquivalenzklassen: [0, 24] für eine Frequenz, die »tief«, und [25, 70] für eine, die »hoch« ist.
- Der Wertebereich von *Umsatz* befindet sich innerhalb von [1, 15.000] € und ist in zwei Äquivalenzklassen unterteilt: [1, 1099] € gelten als »schwach«, und ein Umsatz über 1100 € als »stark«.

Die Definition dieser vier Äquivalenzklassen bzw. Terme zieht eine zweidimensionale *Matrix* auf (Abb. 4a). Mit jeder Klasse können eine Punktzahl sowie Strategien des Onlinemarketings verknüpft werden: Hat ein Kunde eine hohe Kauffrequenz und einen starken Umsatz (Klasse C1; 60 + 60 = 120 Punkte), sollte der Kunde durch ein Loyalitätsprogramm gebunden werden. Kunden, die öfter, aber nur zu einem geringen Betrag einkaufen (C2; 80 Punkte), sollten zum Beispiel durch Rabatte zu größeren Einkäufen bewegt werden, und diejenigen mit einer tiefen Kauffrequenz (C3; 65 Punkte) sollten öfter auf die Website gelockt werden (z.B. mit Sonderangeboten, Wettbewerben oder Spielen).

Durch das Kontextmodell, das am Research Center FMsquare entwickelt wurde (vgl. [Werro 2008], [Meier et al. 2008]), und durch die Definition einer *Zugehörigkeitsfunktion* μ für jeden Term wird der Klassifikationsraum *unscharf* gemacht

(siehe Abb. 4b). Diese unscharfe Unterteilung hat ein wichtiges Ergebnis: Die *scharfe Klassentrennung verschwindet*, die Übergänge der verschiedenen Klassen werden *fließend*, und in einer unscharfen Klassifikation kann ein Kunde *mehreren Klassen gleichzeitig* angehören, wobei sein Zugehörigkeitsgrad zu jeder Klasse exakt berechnet werden kann.

Trennscharf betrachtet (vgl. Abb. 4a) gehört im Beispiel Huber derselben Klasse (C4) an wie Schneider, sie erhalten 25 Punkte. Das ist aber fahrlässig, denn Huber hat fast die gleichen Umsatz- und Frequency-Werte wie Weber, der mit 120 Punkten zu 480 % höher bewertet ist. Wenn das CRM nun der Strategie dieser Klasse folgt, *minimiert* es fälschlicherweise das Marketing in Huber. Key Account Müller andererseits ärgert sich womöglich, dass er vom Unternehmen gleich behandelt wird wie der gleich klassifizierte Weber, dessen Kauffrequenz und Online-Umsatz viel tiefer ist.

Mit *unscharfer Klassifikation* (Abb. 4b) werden solche Fehler vermieden. Huber gehört jetzt teilweise zu *allen Klassen*, mit folgenden *Zugehörigkeitsanteilen*: C1: 20 % ($0,42 \times 0,47$), C2: 27 % ($0,58 \times 0,47$), C3: 22 % ($0,42 \times 0,53$), C4: 31 % ($0,58 \times 0,53$); dies sind Beispielswerte (für die Formel zur Berechnung siehe Abschnitt 4.2). Auch Weber gehört unscharf anteilmäßig zu vier Klassen gleichzeitig. Für Müller (100 % C1) und Schneider (100 % C4) ändert sich aufgrund ihrer extremen Positionen nichts.

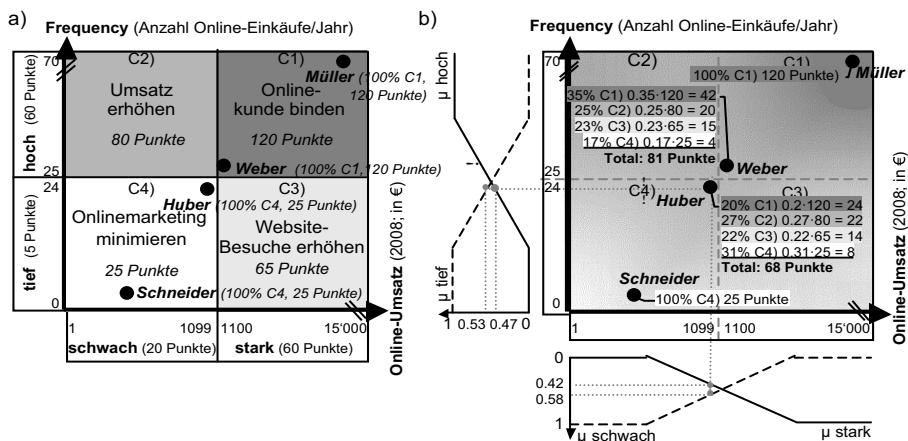


Abb. 4: (a) Scharf und (b) unscharf klassifiziertes Online-Kundenportfolio

Dank den Zugehörigkeitsfunktionen und den Zugehörigkeitsanteilen können die Punkte genau berechnet werden (vgl. Abb. 4b): Huber kriegt mit 68 Punkten unscharf deutlich mehr Punkte als scharf (25), Weber mit 81 hingegen weniger (vorher 120). Die unscharfe Klassifikation erlaubt also auch eine genauere Kalkulation der Scoring-Punkte.

3.4 Fuzzy eRFM Data Cube

Betrachtet man nun zusätzlich die dritte Dimension *Recency* (Kaufaktualität), ergeben sich im *Fuzzy-RFM-Modell* acht scharfe Klassen (vgl. Abb. 5a). Durch die Definition von Fuzzy Sets und Zugehörigkeitsfunktionen wird der

dreidimensionale Klassifikationsraum unscharf (vgl. Abb. 5b).

Dank der unscharfen Klassifikation kann die Punkteanzahl wiederum genau berechnet werden: Unscharf erhält Huber 43,3 Punkte, also deutlich mehr als die 30 Punkte in der scharfen Klassifikation. Auch Kunde Schneider erhält mehr Punkte (scharf: 0, unscharf: 12,9), Weber mit 51,8 Punkten erhält hingegen fast die Hälfte weniger als in der scharfen Klassifikation.

In einer *multidimensionalen Klassifikation* kann man beliebig viele (also n) Variablen hinzufügen und die Kunden mit einem *Fuzzy-Scoring-Modell* anhand verschiedener Kriterien flexibel und differenziert bewerten.

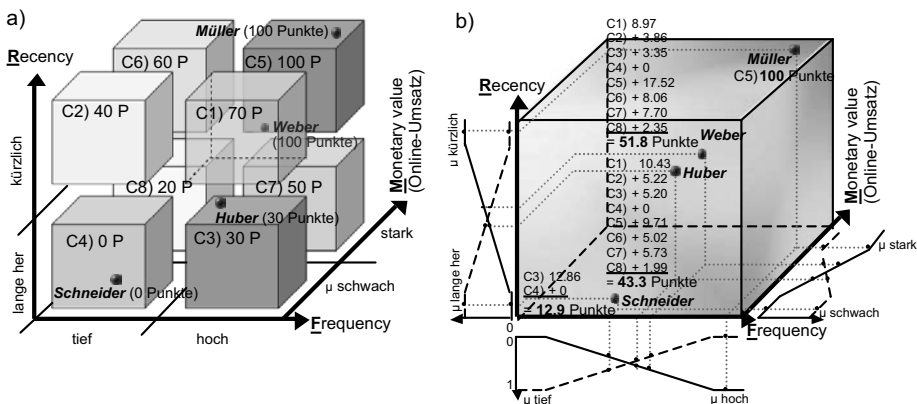


Abb. 5: (a) Sharp & (b) Fuzzy eRFM Data Cube

4 Fallstudie coop@home

4.1 Daten der Fallstudie

In diesem Kapitel soll die unscharfe Klassifikation sowie die RFM-Methode an einem Beispiel mit realen Daten demonstriert werden. Es wurde eine Fallstudie mit Coop, einem führenden Schweizer Detailhändler, durchgeführt. Coop betreibt seit dem Jahr 2000 die Internetplattform coop@home (www.coopathome.ch), auf der die Kunden aus einem großen Sortiment des Online-Produktkatalogs Waren online bestellen können. Zurzeit gibt es rund 12.000 Artikel, die online angeboten und offline nach Hause geliefert werden. Der Kundenstamm von coop@home beträgt insgesamt circa 250.000 registrierte Kunden. Von diesen Kunden haben jedoch mehr als die Hälfte im letzten Jahr oder länger nicht mehr bestellt. Die Selektion der Daten wurde deshalb eingeschränkt auf die Kunden, die im Jahr 2008 mindestens einmal bestellt haben. Aus dieser Menge wurde eine Zufallsstichprobe von 500 Kunden genommen und deren Daten zu den Kennzahlen *Online-Umsatz* (Monetary Value) und *Anzahl Online-Einkäufe* im Jahr 2008 berechnet. Im Durchschnitt hat jeder Kunde 4,6-mal online bestellt und dabei 237 CHF pro Einkauf ausgegeben. Das ergibt einen durchschnittlichen Umsatz von 1100 CHF pro Kunde für das Jahr 2008. Die Höchstwerte liegen bei 14.500 CHF Umsatz und 61 Bestellungen, die Tiefstwerte bei 1 Bestellung und 1 CHF Umsatz.

4.2 Berechnung der Fuzzy-Werte

Bei der *Normalisierung* wird für die Kennzahlen Online-Umsatz und Frequency für jeden Kunden ein Wert zwischen null und eins berechnet, entsprechend seinem jeweiligen Umsatz und der Anzahl seiner Bestellungen. Um diese Be-

rechnung durchführen zu können, muss eine obere Grenze für die Werte Umsatz und Frequency definiert werden. Da der Wert des Kunden mit dem höchsten Umsatz 14.500 ist, kann als *Obergrenze* zum Beispiel 15.000 CHF gewählt werden. Ein Kunde mit einem Umsatz von 10.500 hat mit einem Wert von 0,7 ($1 / 15.000 \times 10.500$) einen starken Umsatz, und mit dem Gegenwert von 0,3 einen schwachen Umsatz. Da der Höchstwert bei der Einkaufsfrequenz bei 61 Bestellungen liegt, wurde als obere Grenze ein Wert von 65 genommen.

Die Verteilung der Daten verläuft exponentiell und hat eine große Varianz (vgl. Abb. 6b/c). Das heißt, die Spitzenwerte übersteigen die Durchschnittswerte von 1100 CHF Umsatz und 4,6 Bestellungen deutlich.

Im Vergleich zu den Höchstwerten würde der Großteil der Kunden schlecht abschneiden und entsprechend zu tief klassifiziert werden, wenn die Klassifikation linear zwischen der Unter- und Obergrenze erfolgen würde. Der Wertebereich wurde deshalb so angepasst, dass nur die Kunden bis maximal 4000 CHF Umsatz und bis höchstens 15 Bestellungen berücksichtigt werden. Das führt dazu, dass Kunden, die diese Werte überschreiten, in der Klassifikation nicht berücksichtigt bzw. nicht mehr dargestellt werden. Da diese Kunden jedoch eindeutig in die Klasse C1 gehören, unabhängig davon, ob die Klassifikation scharf oder unscharf erfolgt, hat dies keine weiteren Konsequenzen.

Für eine unscharfe Klassifikation sind vor allem die Bereiche interessant, in denen viele Kunden (mit ähnlichem Kaufverhalten) *nahe beisammen* liegen und durch scharfe Grenzen fälschlicherweise in eine andere Klasse eingeteilt werden (Kunde 1 und 2 in Abb. 6a).

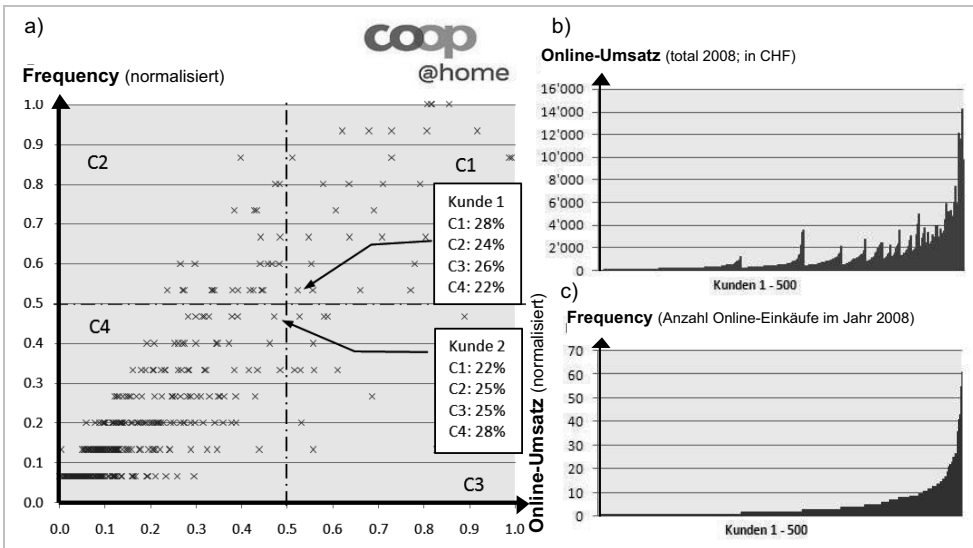


Abb. 6: (a) Unschärfe Klassifikation, (b) Umsatz und (c) Kauffrequenz von coop@home-Kunden

Mit dem *Gamma-Operator* (dem »Kompensatorischen Und«),

$$\mu_{A_i}(x) = \left(\prod_{i=1}^m \mu_i(x) \right)^{(1-\gamma)} \left(1 - \prod_{i=1}^m (1 - \mu_i(x)) \right)^{\gamma}, \quad x \in X, 0 \leq \gamma \leq 1$$

der von [Zimmermann 1992] empirisch getestet und validiert wurde, können die *Zugehörigkeitsanteile* jedes Kunden zu jeder Klasse berechnet werden. Der Kunde 1 (in Abb. 6a) z.B. gehört zu

- 28 % = $((0,56 \cdot 0,53)^{1-5} ((1-0,56) \cdot (1-0,53)^5) / 1,73))$ zu C1
- 24 % = $((0,44 \cdot 0,53)^{1-5} ((1-0,44) \cdot (1-0,53)^5) / 1,73))$ zu C2
- 26 % = $((0,56 \cdot 0,47)^{1-5} ((1-0,56) \cdot (1-0,47)^5) / 1,73))$ zu C3
- 22 % = $((0,44 \cdot 0,47)^{1-5} ((1-0,44) \cdot (1-0,47)^5) / 1,73))$ zu C4

Kunde 1 gehört also genauso wie Kunde 2 zu allen vier Klassen gleichzeitig – mit einem bestimmten Zugehörigkeitsanteil. Anhand der

Daten in Abbildung 6a ist gut zu erkennen, dass die Daten von der schlechtesten Klasse C4 (links unten) hin zur besten Klasse C1 (rechts oben) relativ gleichmäßig durch die Mitte verlaufen.

Es gibt relativ wenige Kunden, die eindeutig zu der Klasse C2 oder C3 gehören, d.h., nur bei wenigen Kunden ist eine der Variablen (Umsatz oder Frequenz) sehr deutlich ausgeprägt. Die beiden Variablen Umsatz und Frequenz korrelieren ziemlich stark, d.h., ein Kunde mit vielen Bestellungen hat meistens auch einen höheren Umsatz.

4.3 Vergleich scharfe und unscharfe Klassifikation

Werden in der scharfen Klassifikation der Coop-Onlinekunden nur diejenigen mit einem starken Umsatz und einer hohen Kauffrequenz berücksichtigt, erhält man die scharfe Klasse C1 (Abb. 7a).

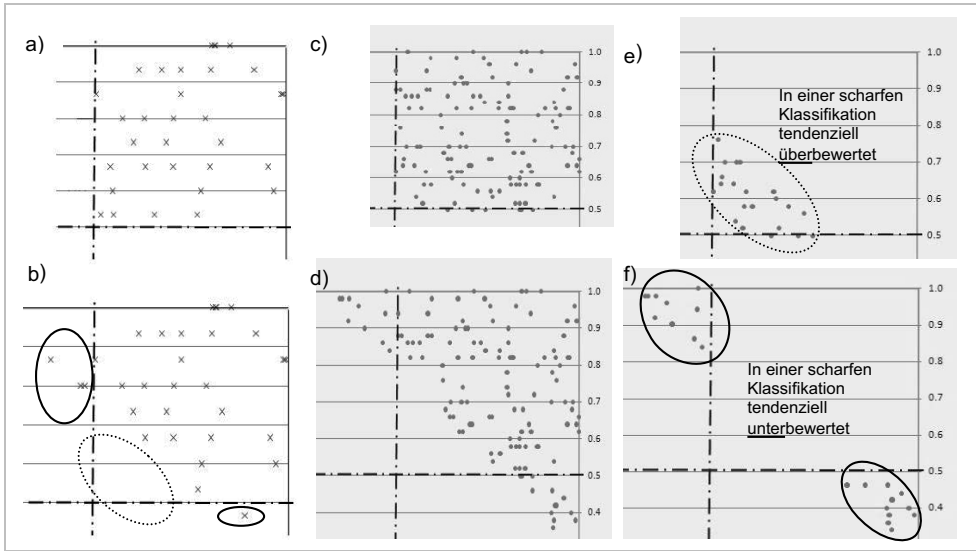


Abb. 7: (a) Scharfe und (b) unscharfe Klassifikation von C1, Fallstudie Coop
(c) Scharfe und (d) unscharfe Klassifikation von C1, bei Gleichverteilung
(e) 19 % der Kunden aus C1, die in der unscharfen Klassifikation herausfallen
(f) 19 % der Kunden aus C2 und C3, die in der unscharfen Klassifikation berücksichtigt werden

In der unscharfen Klassifikation, können diejenigen Kunden ausgewählt werden, die z.B. zu mindestens 34 % zu der Klasse C1 angehören. Im Vergleich zur scharfen Klassifikation werden nun zusätzliche Kunden aus den Klassen C2 und C3 selektioniert, andere aus C1 fallen jedoch weg (Abb. 7b).

In einem Beispiel mit einer Gleichverteilung können die Unterschiede zwischen der scharfen und unscharfen Klassifikation noch deutlicher aufgezeigt werden. Deshalb wurden Zufallswerte für 500 Beispielkunden generiert und diejenigen ausgewählt, die zur Klasse C1 gehören. Dies sind in der scharfen Klassifikation 131 Kunden (siehe Abb. 7c).

Wird die gleiche Anzahl an »besten« Kunden unscharf selektioniert, ist die Verteilung im Vergleich zur scharfen Klassifikation eine andere. 25 Kunden (19 %), die trennscharf betrachtet zu C1 gehören, *fallen* in der unscharfen Klassifikation *heraus* (Abb. 7e). Stattdessen werden 25 Kunden aus den Klassen C2 und C3 berücksich-

tigt, die unscharf ebenfalls zu Klasse C1 gehören (Abb. 7f).

Das heißt, dass bei der *Kundenwertbetrachtung* in einer scharfen Klassifikation die Kunden in Abbildung 7e tendenziell *über-* und in Abbildung 7f *unterbewertet* sind. Bei einer Kundensegmentierung empfiehlt sich, diese Grenzfälle genauer zu analysieren, da hier das *Optimierungspotenzial* am höchsten scheint.

5 Unscharfe Klassifikation weiterer Web- und eCommerce-Kennzahlen

Ein Besucher kann auf unterschiedliche Art auf eine Website wie www.coopathome.ch gelangen: Entweder gibt er die URL (Uniform Resource Locator) direkt in die Adresszeile des Browsers ein oder er hat ein *Lesezeichen* auf die Seite gesetzt oder er gelangt über einen *Verweis* (Hyperlink) einer Drittseite auf die Website (vgl. Abb. 8). Verweise stammen meist von *Suchmaschinen* wie z.B. Google, können aber auch unbezahlte oder bezahlte *Links* (Bannerwerbung)

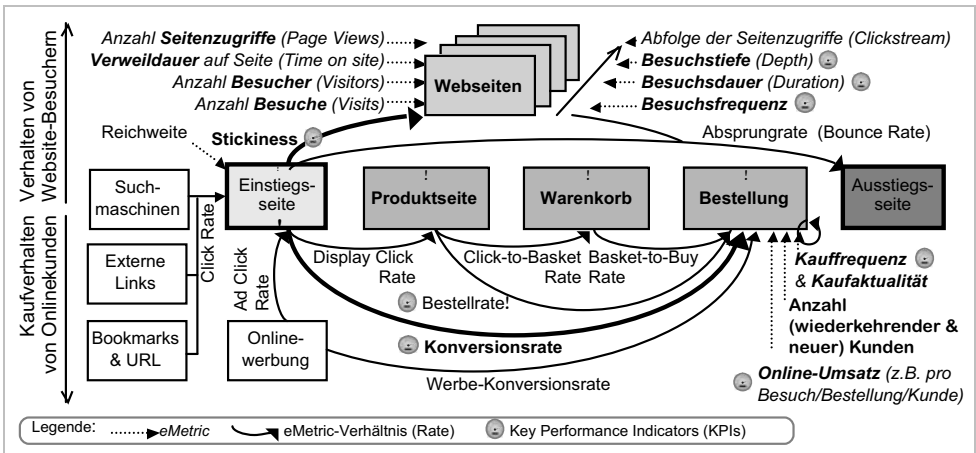


Abb. 8: Die wichtigen Web- und eCommerce-Kennzahlen im Überblick

sein. Ein Besucher verlässt die *Einstiegsseite* (Landing page) wieder, wenn ihn die gebotenen Inhalte nicht interessieren (dies definiert die sogenannte *Absprungrate* bzw. Bounce Rate), oder aber er klickt sich durch weitere Seiten. In diesem Zusammenhang spricht man von der *Page Stickiness* (Website-Haftung).

Wie sich Besucher auf der Website verhalten, wird anhand verschiedener Webkennzahlen analysiert. Zu den wichtigsten gehören die Anzahl *Seitenzugriffe* (Page Views bzw. Page Impressions), *Besucher* (Visitors, Users), *Besuche* (Visits, Sessions), *Besuchstiefe* (Anzahl Seiten pro Besuch) und die *Verweildauer*.

Verfügt die Website über einen Onlineshop, ist die Analyse des Klick- und Kaufverhaltens der Benutzer besonders aufschlussreich. Welche Produkte oder Dienstleistungen haben sich die (potenziellen) Kunden angeschaut (dies definiert aggregiert die *Display Click Rate*)? Welche Produkte haben sie z.B. besonders oft oder nur ganz selten in den Warenkorb gelegt (*Click-to-Basket Rate*) und welche dann auch bestellt (*Basket-to-Buy Rate*)? Die Antworten und Gründe auf solche Fragen sind von strategischer Bedeutung, sei es in Hinblick auf die *Optimierung*

der Website, des eShops, des eMarketings (z.B. Kampagnen, Newsletter, Suchmaschinenoptimierung) und vor allem der Sortimentspolitik. Die *Konversionsrate*, d.h. der Prozentsatz an Besuchern, der zu Käufern konvertiert, und ihre Unterformen gehören zu den wichtigen *Key Performance Indicators* (KPIs) des Webcontrollings [Hukemann 2004; Conrady 2006]. Die Anzahl *Bestellungen* neuer und wiederkehrender Kunden, der durchschnittliche *Bestellwert*, die *Bestellrate* und der *Online-Umsatz* (gemessen in einer Periode oder pro Besuch, Bestellung, Besucher oder pro Kunde) sind weitere kritische Größen des Online-Erfolgs.

Alle Metriken haben gemeinsam, dass man sie oft klassifiziert und segmentiert: Die Ausprägungsgröße einer Webkennzahl ist (eher) tief, mittel oder (eher) hoch, sei es im Soll-Ist-Vergleich für die strategische Planung, im Zeitvergleich zum Vortag, -monat, -quartal und -jahr oder im Konkurrenzvergleich. Wie erläutert ist eine solch scharfe Klassifikation problematisch. Deshalb empfiehlt es sich, alle Kennzahlen unscharf zu klassifizieren, insbesondere wenn sie an weitreichende Maßnahmen und Managemententscheidungen gebunden sind.

6 Nutzenbetrachtung

Die scharfe Klassifikation birgt gewisse Risiken in sich, deshalb schlägt dieser Beitrag einen unscharfen Klassifikationsansatz zur Segmentierung von Web- und eCommerce-Kennzahlen vor. Dieser erlaubt

- die *Reduktion von Komplexität* bei großen Datenmengen, zugleich
- eine *genaue Klassifikation* deren Werte,
- die *exakte Segmentierung* des Traffics, von Besuchern und Onlinekunden,
- eine Reduktion von *Fehlklassifikationen*,
- die Verwendung von *natürlichen Begriffen* des eMarketings und von *linguistischen Variablen* wie z.B. »ziemlich hoher«, »mittelmäßiger« oder »sehr hoher« Umsatz [Meier & Stormer 2008],
- die Berechnung von *individuellen Anreizen* (z.B. Rabatte, Gutscheine, Geschenke),
- die Realisierung von *Mass Customization* [Piller 2006] und *One-to-One-Marketing* und die
- Optimierung des *eBusiness-Erfolgs* durch
- das *Ausschöpfen von Kundenpotenzialen*.

Bei der unscharfen Klassifikation treten auch Nachteile und Probleme auf, z.B. die interne und externe *Kommunikation* dieses Segmentierungsansatzes oder die Herleitung und Definition der *Zugehörigkeitsfunktionen*. Mit dem Unschärfe-Ansatz löst man zwar das Klassifikationsproblem, aber wie im eMarketing strategisch und operativ mit unscharfen Kundensegmenten konkret umgegangen werden soll, ist eine andere Frage und kann nur im betrieblichen, individuellen Kontext beantwortet werden.

7 Literatur

- [Anderson et al. 2004] *Anderson, E.; Fornell, C.; Mazvanheryl, S.*: Customer Satisfaction and Shareholder Value. In: Journal of Marketing, Vol. 68, Oktober 2004, S. 172-185.
- [Bruhn & Homburg 2008] *Bruhn, M.; Homburg, Ch.*: Handbuch Kundenbindungsmanagement. 4. Aufl., Gabler, Wiesbaden, 2008.
- [Conrady 2006] *Conrady, R.*: Controlling des Internet-Auftritts. In: Reinecke, S.; Tomczak, T.: Handbuch Marketing-Controlling. 2. Aufl., Gabler, Wiesbaden, 2008.
- [Günter & Helm 2006] *Günter, B.; Helm, S. (Hrsg.)*: Kundenwert. Gabler, Wiesbaden, 2006.
- [Hassler 2008] *Hassler, M.*: Web Analytics. mitp, Heidelberg, 2008.
- [Hukemann 2004] *Hukemann, A.*: Controlling im Onlinehandel – Ein kennzahlenorientierter Ansatz für Onlineshops. Logos, Berlin, 2004.
- [Liu 2007] *Liu, B.*: Web Data Mining – Exploring Hyperlinks, Contents, and Usage Data. Springer-Verlag, München, 2007.
- [Krafft 2007] *Krafft, M.*: Kundenbindung und Kundenwert. 2. Aufl., Physica, Heidelberg, 2007.
- [McCarthy & Hastak 2006] *McCarthy, J.; Hastak, M.*: Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. In: Journal of Business Research, Vol. 60, Juni 2006, S. 656-662.
- [Meier & Stormer 2008] *Meier, A.; Stormer, H.*: eBusiness & eCommerce. Springer-Verlag, 2008.
- [Meier et al. 2008] *Meier, A.; Schindler, G.; Werro, N.*: Fuzzy Classification on Relational Databases. In: Galindo, J. (Hrsg.): Handbook of Research on Fuzzy Information Processing in Databases. Idea Group, Hershey, 2008.
- [Neckel & Knobloch 2005] *Neckel, P.; Knobloch, B.*: Customer Relationship Analytics. dpunkt.verlag, Heidelberg, 2005.
- [Piller 2006] *Piller, F.*: Mass Customization: Ein wettbewerbsstrategisches Konzept im Informationszeitalter. Gabler, Wiesbaden, 2006.

- [Werro 2008] Werro, N.: Fuzzy Classification of On-line Customers. Dissertation, Universität Fribourg, 2008, <http://ethesis.unifr.ch/theses/downloads.php?file=WerroN.pdf>; Zugriff am 12.01.2009.
- [Wünschmann et al. 2008] Wünschmann, S.; Schwarz, U.; Müller, S. (Hrsg.): Website-Gestaltung. Heidelberg, 2008.
- [Zadeh 1965] Zadeh, L.: Fuzzy Sets. In: Information and Control, Vol. 8, 1965, S. 338-353.
- [Zimmermann 1992] Zimmermann, H.-J.: Fuzzy Set Theory – and its Applications. 2. Aufl., Kluwer Academic Publishers, London, 1992.

Darius Zumstein MA
Patrik Hugli BSc
Universität Fribourg
Information Systems Research Group
Boulevard de Pérolles 90
CH-1700 Fribourg
{darius.zumstein,
patrik.hugli}@unifr.ch
<http://diuf.unifr.ch/is>



Ralf Kneuper, Ernest Wallmüller
(Hrsg.)

CMMI in der Praxis

Fallstudien zur Verbesserung der
Entwicklungsprozesse mit CMMI

2009, 221 Seiten, Festeinband

€ 39,00 (D)

ISBN 978-3-89864-571-3



dpunkt.verlag

Ringstraße 19 B · D-69115 Heidelberg · fon: 0 62 21 / 14 83 40

fax: 0 62 21 / 14 83 99 · e-mail: bestellung@dpunkt.de · www.dpunkt.de